CHAPTER 3

Brownian Motion and Its Applications

1. Definition of Brownian Motion

In the next Chapter we will provide a reasonably systematic introduction to stochastic processes; we start however by considering a particular stochastic process that is of particular importance both in theory and in the applications, and then consider some of its uses.

DEFINITION. A stochastic process (in the strict sense) is a function $v(\omega, t)$ of two arguments where $(\Omega, \mathcal{B}, P)$ is a probability space, $\omega \in \Omega$, and $t \in \mathbb{R}$, such that for each ω $v(\omega, t)$ is a function of t and for each t $v(\omega, t)$ is a random variable.

If t is a space variable the stochastic process is usually called a random field.

DEFINITION. "Brownian Motion" (in mathematical terminology) is a stochastic process $w(\omega, t)$, $\omega \in \Omega$, $0 < t < 1$, that satisfies the following four axioms:

(1) $w(\omega, 0) = 0$ for all ω.
(2) For each ω, $w(\omega, t)$ is a continuous function of t.
(3) For each t, $w(\omega, t)$ is a Gaussian variable with mean zero and variance t.
(4) $w(\omega, t)$ has independent increments, i.e., if $t_1 < t_2 \leq t_3 < t_4$ then the random variables $w(\omega, t_2) - w(\omega, t_1)$ and $w(\omega, t_4) - w(\omega, t_3)$ are independent.

Note that what is called in mathematics Brownian Motion (BM) is called in physics the Wiener process. And what is called in physics BM is a different process which is called in mathematics the Ornstein-Uhlenbeck process, which we shall discuss a little bit later.

First of all one must show that a process that satisfies all of these conditions exists. This is not a trivial issue; we shall see shortly that if the second condition above is replaced by the requirement that w be differentiable, then there is no way to satisfy the conditions. The

original proof of Wiener consisted of showing that the Fourier series

$$\frac{\pi}{2\sqrt{2}} \sum_{k=1}^{\infty} \frac{a_k}{k} \sin(\pi k t/2),$$

where the a_k are independent Gaussian variables with mean zero and variance one converges, and its sum satisfies the conditions above for $0 \leq t \leq 1$. Each coefficient is a random function defined on some probability space $(\Omega, \mathcal{B}, P)$, and the resulting BM is also a function on the very same Ω. For longer times one can construct the process by stringing the processes constructed by this series end to end. We refer the reader to the literature.

Next we derive some consequences of the definition of BM.

(1) $w(\omega, t_2) - w(\omega, t_1)$, where $t_2 > t_1$, is a Gaussian variable with mean zero and variance $t_2 - t_1$. Indeed, the Gaussian variable $w(\omega, t_2)$, which has mean zero and variance t_2 by definition, may be written as a sum of two Gaussian variables

$$w(\omega, t_2) = [w(\omega, t_2) - w(\omega, t_1)] + [w(\omega, t_1) - w(\omega, 0)].$$

By axiom 4, those two variables are independent and hence the variance of their sum (which is t_2) is the sum of their variances. The variance of $[w(\omega, t_1) - w(\omega, 0)]$ is t_1, and therefore the variance of $[w(\omega, t_2) - w(\omega, t_1)]$ is $t_2 - t_1$. One can also check that any linear combination of Gaussian variables is Gaussian.

(2) The correlation function of BM is $E[w(t_1)w(t_2)] = \min\{t_1, t_2\}$. Indeed, assuming $t_1 < t_2$ we get

$$\begin{aligned} E[w(t_1)w(t_2)] &= E[w(t_1)(w(t_1) + (w(t_2) - w(t_1))] \\ &= E[w(t_1)w(t_1)] + E[w(t_1)(w(t_2) - w(t_1))] = t_1. \end{aligned}$$

In the last equation, the variables $w(t_1)$ and $w(t_2) - w(t_1)$ are independent and each has mean zero.

(3) Consider the variable

$$\frac{w(\omega, t + \Delta t) - w(\omega, t)}{\Delta t}.$$

It is Gaussian with mean zero and variance $(\Delta t)^{-1}$, which tends to infinity as Δt tends to zero. Therefore one can guess that the derivative of $w(\omega, t)$ for any fixed ω exists nowhere with probability one.

Non-differentiable functions may have derivatives in the sense of distributions. The derivative in the sense of distributions $v(\omega, s)$ of a

BM $w(\omega, t)$ is called "white noise" and is defined by the property:

$$\int_{t_1}^{t_2} v(\omega, s)ds = w(\omega, t_2) - w(\omega, t_1).$$

The origin of the name will be clarified in the next chapter.

Two-dimensional BM is $(v_1(\omega, t), v_2(\omega, t))$ where v_1, v_2 are independent BM's, and similarly for n-dimensional BM.

2. A Relation Between Brownian Motion and the Heat Equation

We first solve the heat equation

$$v_t = \frac{1}{2}v_{xx}, \quad v(x, 0) = \phi(x) \tag{3.1}$$

on $-\infty < x < \infty$, $t > 0$, by Fourier transforms. We start with the following observations. Let $g = g(x, t)$ and $\hat{g}(k, t)$ be its Fourier transform in x

$$g(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx}\hat{g}(k, t)dk.$$

Then

$$\frac{\partial g}{\partial x} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ike^{ikx}\hat{g}(k, t)dk$$

and taking the transform of both sides yields

$$\widehat{\frac{\partial g}{\partial x}} = ik\hat{g}(k).$$

Similarly

$$\frac{\partial g}{\partial t} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx}\frac{\partial \hat{g}}{\partial t}(k)dk,$$

and

$$\widehat{\frac{\partial g}{\partial t}} = \frac{\partial \hat{g}}{\partial t}.$$

Now return to the heat equation. Taking the Fourier transform with respect to x gives

$$\hat{v}_t(k, t) = -\frac{1}{2}k^2\hat{v}(k, t), \quad \hat{v}(k, 0) = \hat{\phi}(k).$$

Which has as solution

$$\hat{v}(k, t) = Ce^{-k^2t/2}.$$

Applying the inverse transform to $\hat{v}(k, t)$ and taking into account that the inverse transform of $e^{-k^2t/2}$ is $t^{-1/2}e^{-x^2/2t}$, the inverse transform of

$\hat{\phi(k)}$ is $\phi(x)$, and the inverse transform of a product is a convolution, we get

$$v(x,t) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-(x-x')^2/2t}\, \phi(x')dx'. \tag{3.2}$$

The function

$$G(x) = \frac{e^{-x^2/2t}}{\sqrt{2\pi t}}$$

is the Green function of the heat equation and we have shown that the solution of the heat equation is the convolution of the initial data with the Green function.

Let η be a Gaussian variable with mean x and variance t. Using (2.4) shows that

$$v(x,t) = E[\phi(\eta)].$$

Remember that $w(t)$, where w is BM, is a Gaussian variable with mean zero and variance t. Start a BM at the point (x,t). Let $w(\omega,t)$ be a BM going backwards from (x,t). It will reach the x-axis after a time interval t at the point $\eta = x + w(\omega,t)$ where the initial value is $\phi(\eta)$. The expected value of the function $\phi(\eta)$ is, as we have just seen, $v(x,t)$. Thus we can evaluate the solution $v(x,t)$ of the heat equation by starting BM's from (x,t) and averaging the initial data at the points where these BM's hit the x-axis.

3. Solution of the Heat Equation by Random Walks

We now rederive the result above in a roundabout way which will be useful to our analysis of a more general situation. We will construct a grid to approximate the heat equation (3.1), solve the resulting discrete equations by a random walk, and take a limit that will reproduce the result of the previous section. To construct the grid, draw horizontal and vertical lines in the (x,t) plane. The distance between the horizontal lines is k (*not* the Fourier variable!) and between the vertical lines is h. The points at which these lines intersect will carry values of an approximation V of the solution $v(x,t)$ of the heat equation. That is, each point $(i,n) = (ih, nk)$ on the grid carries a value of the grid function

$$V_i^n = V(ih, nk).$$

Construct a difference approximation of the derivatives in (3.1)

$$v_t = \frac{V_i^{n+1} - V_i^n}{k} + O(k) \tag{3.3}$$

$$v_{xx} = \frac{V_{i+1}^n + V_{i-1}^n - 2V_i^n}{h^2} + O(h^2). \tag{3.4}$$

Substituting (3.3) and (3.4) into (3.1) we obtain an equation for the V_i^n

$$\frac{V_i^{n+1} - V_i^n}{k} = \frac{1}{2}\,\frac{V_{i+1}^n + V_{i-1}^n - 2V_i^n}{h^2}. \tag{3.5}$$

Starting from the initial data $V_i^0 = \phi(ih)$ we can find a solution of (3.5) at time $t = nk$ for any n by the recurrence formula

$$V_i^{n+1} = V_i^n + \lambda(V_{i+1}^n + V_{i-1}^n - 2V_i^n) = (1-2\lambda)V_i^n + \lambda V_{i+1}^n + \lambda V_{i-1}^n, \tag{3.6}$$

where

$$\lambda = \frac{1}{2}\frac{k}{h^2}.$$

Define the local "truncation error"

$$\tau_i^n = \frac{v_i^{n+1} - v_i^n}{k} - \frac{1}{2}\frac{v_{i+1}^n + v_{i-1}^n - 2v_i^n}{h^2} - (v_t - v_{xx}),$$

where v is a smooth function and $v_i^n = v(ih, nk)$. From (3.3) and (3.4) one finds that that

$$\tau_i^n = O(k) + O(h^2).$$

In numerical analysis the fact that τ_i^n tends to zero as $h \to 0, k \to 0$ is called "consistency." Thus, the scheme (3.5) is consistent and its truncation error is of order $O(k) + O(h^2)$.

Now we show that for $\lambda \le 1/2$ the approximate solution V converges to the exact solution v as h and k tend to zero. It is easy to check that the error $e_i^n = v_i^n - V_i^n$ satisfies the equation

$$e_i^{n+1} = (1-2\lambda)e_i^n + \lambda e_{i+1}^n + \lambda e_{i-1}^n + k\tau_i^n.$$

Taking the absolute value of both sides we get

$$|e_i^{n+1}| \le (1-2\lambda)|e_i^n| + \lambda|e_{i+1}^n| + \lambda|e_{i-1}^n| + k|\tau_i^n|$$

where we have assumed that $1 - 2\lambda \ge 0$ (or $\lambda \le 1/2$). Define

$$E^n = \max_i |e_i^n|, \tag{3.7}$$

and let

$$\tau^n = \max_i |\tau_i^n|, \quad \tau = \max_n |\tau^n|. \tag{3.8}$$

Then

$$E^{n+1} \le E^n + k\tau^n \le E^n + k\tau$$

thus

$$E^{n+1} \le E^n + k\tau \le E^{n-1} + 2k\tau \le \ldots \le E^0 + (n+1)k\tau.$$

If we start from the exact solution then $E^0 = 0$ and hence

$$E^n \le nk\tau = t\tau.$$

Recall that the local truncation error tends to zero as $h, k \to 0$ and consider the solution of the heat equation on a finite t interval $0 \leq t \leq T$ for some finite T; then E^n tends to zero as h and k tend to zero provided $\lambda = k/2h^2$ is less than or equal to $1/2$. That means that the approximate solution coverges to the exact solution for $\lambda \leq 1/2$.

Choose $\lambda = 1/2$. Then (3.6) becomes

$$V_i^{n+1} = \frac{1}{2}(V_{i+1}^n + V_{i-1}^n). \tag{3.9}$$

Using (3.9) and iterating backwards in time we can write w_i^n in terms $w_i^0 = \phi(ih)$

$$\begin{aligned} V_i^n &= \frac{1}{2}V_{i+1}^{n-1} + \frac{1}{2}V_{i-1}^{n-1} \\ &= \frac{1}{4}V_{i-2}^{n-2} + \frac{2}{4}V_i^{n-2} + \frac{1}{4}V_{i+2}^{n-2} \\ &= \dots \\ &= \sum_{j=0}^{n} C_{j,n}\phi((-n+2j+i)h). \end{aligned}$$

It is easy to see that the numbers $C_{j,n}$ are the binomial coefficients divided by 2^n

$$C_{j,n} = \frac{1}{2^n}\binom{n}{j}. \tag{3.10}$$

Thus

$$V_i^n = \sum_{j=0}^{n} \frac{1}{2^n}\binom{n}{j}\phi((-n+2j+i)h). \tag{3.11}$$

We can interpret the numbers $C_{j,n}$ as follows: Imagine that a drunkard makes a step h to the left with probability $1/2$ or a step h to the right with probability $1/2$ starting from the point (i, n) (see Figure 1) For the moment ignore the fact that the walk is backwards in time. Assume that her successive steps are independent. The probability that she will reach the point $(i-j, 0)$ after n steps is exactly $C_{j,n}$. We can represent this drunken walk as a sum of n random variables

$$\eta_k = \begin{cases} h, & \text{probability } \frac{1}{2} \\ -h, & \text{probability } \frac{1}{2} \end{cases}, \tag{3.12}$$

with $k = 1, 2, \dots n$. This representation gives us another expression for $C_{j,n}$

$$C_{j,n} = P\left(\sum_{k=1}^{n} \eta_k = i - j\right). \tag{3.13}$$

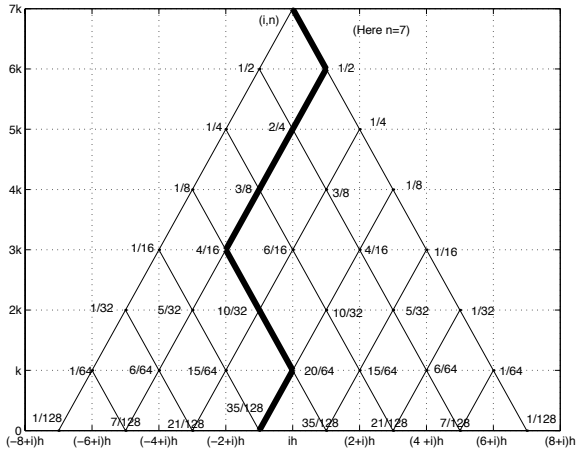

FIGURE 1. Backwards walk for the heat equation.

According to the Central Limit Theorem, the sum $\sum_{k=1}^{n} \eta_k$ tends to a Gaussian variable with mean zero and variance nh^2 as $n \to \infty$. Recall that $\lambda = k/2h^2 = 1/2$, consequently $h^2 = k$ and hence $nh^2 = nk = t$. So $\sum_{k=1}^{n} \eta_k$ tends to a Gaussian variable with mean zero and variance t as $k \to \infty$. Hence,

$$P\left(\sum_{k=1}^{n} \eta_k = i - j\right) \to \frac{e^{-(x-x')^2/2t}}{\sqrt{2\pi t}}.$$

Therefore,

$$w_i^n = \sum_{j=0}^{n} C_{j,n}\phi((-n+2j+i)h) \to \int_{-\infty}^{\infty} \frac{e^{-(x-x')^2/2t}}{\sqrt{2\pi t}}\phi(x')dx' \quad (3.14)$$

as $n \to \infty$. We have in fact used the Central Limit Theorem to derive the solution formula for the heat equation.

4. The Wiener Measure

Having described the properties of Brownian motion (BM), our present goal is to construct a probability measure on the space of continuous functions in such a way that the set of Brownian motions (the samples of BM, the functions $w(\omega, t)$ for various values of ω) have probability 1. In other words, we are trying to define a measure on a space of functions in such a way that the only functions that count are the BM's. To this end, consider the space of continuous functions $u(t)$ such that $u(0) = 0$. This collection is now our sample space Ω. Ω is the space of experimental outcomes and our experiment consist of creating an instance of a continuous function with $u(0) = 0$.

Next we need to define a σ-algebra. Pick an instant in time, say t_1, and associate with this instant a window of values $[a_1, b_1]$. Consider the subset of all the continuous functions that pass through this window and denote it as C_1. This subset is called a cylinder set. For every instant and every window we can define a corresponding cylinder set, i.e., C_i is the subset of all continuous functions that pass through the window $[a_i, b_i]$ at the instant t_i. Next, consider two cylinder sets C_1 and C_2. Then $C_1 \cap C_2$ is the set of functions that pass through both windows. Similarly, $C_1 \cup C_2$ is the set of functions that pass through either of C_1 or C_2. This construction can be carried on to show that the cylinder sets indeed form a σ-algebra on the space of continuous functions in $[0, 1]$ that have the value zero at the origin.

The next step in our construction is to define a measure, i.e., a rule by which to attach probabilities to the cylinder sets. We want to define the measure in such a way that is appropriate for BM's. Take the cylinder set C_1. If the functions that belong to this cylinder set are Brownian motions, the probability of the cylinder set is

$$P(C_1) = \int_{b_1}^{a_1} \frac{e^{-s_1^2/2t_1}}{\sqrt{2\pi t_1}} ds_1.$$

Assign this P to this set, and similarly for other cylinder sets constructed in the same way at different values of t.

Next, consider the intersection $C_1 \cap C_2$ of two cylinder sets C_1 and C_2 with $t_2 > t_1$. By the property of Brownian motion that nonoverlapping increments are independent random variables with Gaussian distributions, we conclude that the probability we should assign to $C_1 \cap C_2$ is

$$P(C_1 \cap C_2) = \int_{b_1}^{a_1} \frac{e^{-s_1^2/2t_1}}{\sqrt{2\pi t_1}} ds_1 \int_{b_2}^{a_2} \frac{e^{-(s_2-s_1)^2/2(t_2-t_1)}}{\sqrt{2\pi(t_2-t_1)}} ds_2.$$

Similarly, we can define a probability for the intersection of a countable number of cylinder sets. To prove that the measure defined in this way is a probability measure we need to show that it satisfies the axioms of probability. For the countable additivity property we can proceed by invoking the duality principle that the complement of the union of some sets is equal to the intersection of their complements, and then use the formula for the probability of the intersection of cylinder sets. The identity $P(\Omega) = 1$ can be seen from the evaluation of the Gaussian integrals in the interval $(-\infty, +\infty)$. The measure we defined is due to Wiener and carries his name.

Suppose that F is a number attached to a continuous function. For example, if $u(s)$ is a continuous function with $u(0) = 0$ and $0 \leq s \leq 1$

then we could define F as $F = \int_0^1 u^2(s)ds$. Any mapping that attaches a number to a function is, for historical reasons, called a functional. Also for historical reasons, a functional acting on a function $u(\cdot)$ is written as $F[u(\cdot)]$. F is a function on Ω, the space of continuous functions that start from the origin.

If one has a measure one has an integral. Denote the integral with respect to the Wiener measure by $\int dW$. In particular, if $\mathcal{X}_C$ is the indicator function of the set C ($\mathcal{X}_C = 1$ if ω is in C, $\mathcal{X}_C = 0$ otherwise), then $\int \mathcal{X}_C dW = P(C)$. If we attach to each BM w a number $F[w(\cdot)]$, (the number is attached to the whole BM), then the integral $\int F[w(\cdot)]dW$ is by definition the expected value of F as w runs over all the possible BM's.

EXAMPLE. Suppose $F[u(\cdot)] = w^2(1)$; i.e., we take a BM w, look at the value of w when $t = 1$, and square that number. This is a number attached to w. $w(1)$ is by definition a Gaussian random variable with mean 0 and variance 1. Then

$$\int F dW = \int_{-\infty}^{+\infty} u^2 \frac{e^{-u^2/2}}{\sqrt{2\pi}} du = 1.$$

EXAMPLE. Fubini's theorem can be extended to integrals more abstract than the elementary finite-dimensional integral and in particular we can show that it is legitimate, under appropriate conditions, to interchange the order of integration with respect to the Wiener measure and ordinary integration. For instance, if $F[w(\cdot)] = \int_0^1 w^2(s)ds$ (a perfectly reasonable way to attach a number to the function $w(t)$), then

$$\int dW \int_0^1 w^2(s)ds = \int_0^1 ds \int dW w^2(s) = \int_0^1 sds = \frac{1}{2}$$

because $w(s)$ is a Gaussian variable with variance s and mean 1.

5. Heat Equation with Potential

Now consider the initial value problem

$$v_t = \frac{1}{2}v_{xx} + U(x)v_x, \quad v(x,0) = \phi(x). \tag{3.15}$$

(Note that with the addition of the imaginary i in front of the time derivative, this would be a Schrödinger equation and U would be a potential.) Generalizing what has been done before, approximate this equation by

$$\frac{V_i^{n+1} - V_i^n}{k} = \frac{1}{2}\frac{V_{i-1}^n + V_{i+1}^n - 2V_i^n}{h^2} + \frac{1}{2}\left(U_{i-1}V_{i-1}^n + U_{i+1}V_{i+1}^n\right) \tag{3.16}$$

where $U_i = U(ih)$ and V_i^n is, as before, a function defined on the nodes (ih, nk) of a grid. Note the clever split of the term Uv into two halves; we now show that the addition of these terms does not destroy the convergence of the approximation to the solution of the differential equation. First check consistency: As before,

$$\frac{v_i^{n+1} - v_i^n}{k} = v_t + O(k), \quad \frac{v_{i+1}^n + v_{i-1}^n - 2v_i^n}{h^2} = v_{xx} + O(h^2).$$

For the potential term we find

$$\begin{aligned}\frac{1}{2}\left(U_{i+1}v_{i+1}^n + U_{i-1}v_{i-1}^n\right) &= \frac{1}{2}\left(2U_i v_i^n + h^2(Uv)_{xx} + h^2 O(h^2)\right) \\ &= V_i v_i^n + O(h^2).\end{aligned}$$

And so

$$\begin{aligned}\frac{v_i^{n+1} - v_i^n}{k} - \frac{1}{2}\frac{v_{i+1}^n + v_{i-1}^n - 2v_i^n}{h^2} - \frac{1}{2}\left(U_{i+1}v_{i+1}^n + U_{i-1}v_{i-1}^n\right) \\ - (v_t - \frac{1}{2}v_{xx} - U(x)v) = O(k) + O(h^2).\end{aligned}$$

Thus the truncation error is small.

Now we show that the approximate solution converges to the exact solution as k and h tend to zero. Let $\lambda = k/2h^2$ as before. The exact solution of (3.15) satisfies

$$v_i^{n+1} = (1-2\lambda)v_i^n + \lambda v_{i+1}^n + \lambda v_{i-1}^n + \frac{k}{2}\left(V_{i+1}v_{i+1}^n + V_{i-1}v_{i-1}^n\right) + \tau_i^n$$

while the approximate solution satisfies

$$V_i^{n+1} = (1-2\lambda)V_i^n + \lambda V_{i+1}^n + \lambda V_{i-1}^n + \frac{k}{2}\left(V_{i+1}w_{i+1}^n + V_{i-1}V_{i-1}^n\right).$$

Thus the error $e_i^n = v_i^n - V_i^n$ satisfies

$$e_i^{n+1} = (1-2\lambda)e_i^n + \lambda e_{i+1}^n + \lambda e_{i-1}^n + \frac{k}{2}(U_{i+1}e_{i+1}^n + U_{i-1}e_{i-1}^n) + k\tau_i^n.$$

Taking the absolute value of both sides and choosing $\lambda \le 1/2$ we get

$$|e_i^{n+1}| \le (1-2\lambda)|e_i^n| + \lambda|e_{i+1}^n| + \lambda|e_{i-1}^n| + \frac{k}{2}(|U_{i+1}||e_{i+1}^n| + |U_{i-1}||e_{i-1}^n|) + k|\tau_i^n|.$$

Assume that the potential is bounded

$$|U(x)| \le M$$

and recall the definitions of E^n (3.7) and τ^n (3.8). It follows that

$$E^{n+1} \le E^n + MkE^n + k\tau^n \le E^n(1+Mk) + k\tau$$

and hence

$$E^{n+1} \le e^{kM}E^n + k\tau.$$

Then

$$\begin{aligned} E^{n+1} &\le e^{kM}E^n + k\tau \\ &\le e^{kM}(e^{kM}E^{n-1} + k\tau) + k\tau \\ &= e^{2kM}E^{n-1} + k\tau(1 + e^{kM}) \\ &\le \dots \\ &\le e^{(n+1)kM}E^0 + k\tau\left(1 + e^{kM} + e^{2kM} + \dots + e^{nkM}\right) \\ &= e^{(n+1)kM}E^0 + k\tau\frac{e^{(n+1)kM} - 1}{e^{kM} - 1}. \end{aligned}$$

Since we start to compute the approximate solution from the given initial condition $v(x, 0) = \phi(x)$, we may assume that $E^0 = 0$. Therefore at time $t = nk$ E^n is bounded by

$$E^n \le k\tau\frac{e^{tM} - 1}{e^{kM} - 1} \le e^{tM}k\tau.$$

And we see that E^n tends to zero as k and h tend to zero with $\lambda \le 1/2$. Thus the approximation is convergent.

Now set $\lambda = 1/2$. Then for the approximate solution we have

$$\begin{aligned} V_i^{n+1} &= \frac{1}{2}(V_{i-1}^n + V_{i+1}^n) + \frac{k}{2}(U_{i+1}V_{i+1}^n + U_{i-1}V_{i-1}^n) \\ &= \frac{1}{2}(1 + kU_{i+1})V_{i+1}^n + \frac{1}{2}(1 + kU_{i-1})V_{i-1}^n. \end{aligned}$$

As before, the approximate solution V may be represented in the form

$$V_i^n = \sum_j C_{ij}^n V_j^0 = \sum_j C_{ij}^n \phi_j.$$

But here, unlike in the case $V = 0$, each movement to the right or to the left brings in not just a factor $1/2$ but also $1/2$ times a factor $(1 + kU(x))$ (see Figure 2).

Thus for the heat equation with potential C_{ij}^n is equal to $1/2$ multiplied by the factor $(1 + kU(hi_j))$ for each step of each path $\{i = i_0, i_1, i_2, \dots, i_n = j\}$ from (i, n) to $(j, 0)$

$$C_{ij}^n = \frac{1}{2^n}(1 + kU(hi_1))(1 + kU(hi_2)) \dots (1 + kU(hi_n))$$

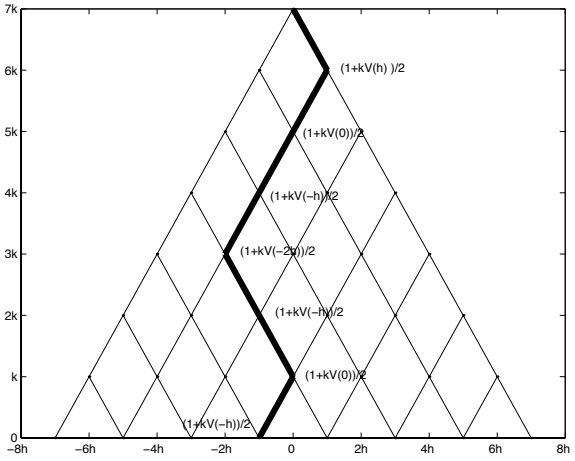

FIGURE 2. Backwards walk for the heat equation with potential.

Since $(1 + kU_i) = e^{kU_i} + O(k^2)$ and each product has $n = O(k^{-1})$ factors, we may rewrite the product as

$$\prod_{\text{path}}(1 + kU_i) = \prod_{\text{path}} \left(e^{kU_i} + O(k)\right)$$
$$= \prod_{\text{path}} e^{kU_i} + O(k) = \exp\left(\sum_{\text{path}} kV_i\right) + O(k).$$

where $\sum_{\text{path}}$ denotes a sum over all the nodes encountered in a single path from i_0 to j; adding up all the paths that contribute to the sum at j we find:

$$V_i^n = \sum_{\text{all paths}} \frac{1}{2^n} e^{\sum_{\text{path}} kV_i}\, \phi_j + O(k).$$

As k and h tend to zero, these paths ("walks") get to resemble BM's, so that

$$\exp\left[\sum_{\text{path}} kV_i\right] \to \exp\left[\int_0^t V(x + w(s))ds\right],$$

where $w(\cdot)$ is a Brownian motion path; $x + w(s)$ is then a BM that starts from x. Taking into account the factor ϕ_j we find the solution of the differential equation in the form

$$v(x,t) = E\left[e^{\int_0^t V(x+u(s))ds} g(x + u(t))\right] = \int dW e^{\int_0^t V(x+u(s))ds} g(x+u(t)). \tag{3.17}$$

This is the Feynman-Kac formula. It reduces to the solution formula for the heat equation when $U = 0$. This result is useful in quantum mechanics and in other fields.

6. Physicists' Notation for Wiener Measure

Physicists use an interesting notation for the Wiener measure which can be useful when one uses Wiener integrals in problems of mechanics and quantum mechanics. There are no new ideas here, just new notation. Before proceeding, we recall a number of results already established.

In the construction of cylinder sets for the Wiener measure, pick an event $C = \bigcap C_i$ where C_i is associated with the interval $[a_i, b_i]$ and $t_i = ih$. Additionally, assume that the windows are of small width, i.e., $b_i - a_i = du_i$. The probability attached to such a set is

$$P = \int_{a_1}^{b_1} \frac{e^{-u_1^2/2h}}{\sqrt{2\pi h}} du_1 \int_{a_2}^{b_2} \frac{e^{-(u_2-u_1)^2/2h}}{\sqrt{2\pi h}} du_2 \cdots \int_{a_n}^{b_n} \frac{e^{-(u_n-u_{n-1})^2/2h}}{\sqrt{2\pi h}} du_n. \tag{3.18}$$

For sufficiently narrow windows, each integral in (3.18) can be approximated by

$$\int_{a_i}^{b_i} \frac{e^{-(u_i-u_{i-1})^2/2h}}{\sqrt{2\pi h}} du_i \approx \frac{e^{-(u_i^*-u_{i-1}^*)^2/2h}}{\sqrt{2\pi h}} du_i,$$

where $u_i^* \in [a_i, b_i]$. Therefore P can be approximated by

$$P \approx \frac{1}{Z} \exp\left(- \sum_{i=1}^{n} \frac{(u_i^* - u_{i-1}^*)^2 h}{2h^2} \right) [du],$$

where $[du] = du_1 du_2 \dots du_n$ and Z is an appropriate normalization constant. Thus, formally (this means "not rigorously" or "not really")

$$dW = \frac{1}{Z} e^{-\frac{1}{2}\int_0^t \left(\frac{du}{ds}\right)^2 ds} [du], \quad Z = \prod_{n\to\infty} (2\pi h)^{n/2}.$$

This expression is formal in the sense that neither the integral in the exponent, the limiting Z, nor the product of du's hidden in $[du]$ exist; we have shown in particular that Brownian motion is not differentiable. But still this expression turns out to be useful.

Recall that, given the equation $v_t = \frac{1}{2} v_{xx} + V(x)v$ with the initial data $v(x, 0) = \phi(x)$, we have

$$v(x,t) = \int e^{\int_0^t G(x+u(s))ds} \phi(x + u(t)) dW.$$

In terms of the new notation, this last integral can be written as

$$v(x,t) = \frac{1}{Z}\int e^{-\int_0^t [\frac{1}{2}\left(\frac{du}{ds}\right)^2 - G(x+u(s))]ds}\phi(x+u(t))[du]. \tag{3.19}$$

By definition (3.19) is a "sum over paths." In principle, one can evaluate it by taking many Brownian motion paths, evaluating the integrals for each path, and then averaging the results. The formal (i.e., meaningless if one looks too closely) expression $[du]$ is often written as "$d$path" (or something similar). Note, and this is an important point, that the exponent is an integral of what we will see is a Lagrangian. Similar integrals appear in quantum mechanics (with an additional imaginary factor i in the exponent).

If one is given an expression for a measure in the form (3.19), one can interpret it properly by retracing the steps that led to that form: the integral of the derivative squared denotes the Wiener measure, the other part of the integral can be discretized, and the terms in the resulting sums become the probabilities of a "path" belonging to a cylinder set.

7. More on the Connection Between Brownian Motion and the Heat Equation

Consider the random variables $w(\omega, t)$ as functions of ω (i.e., as random variables) for several values of t. Define the function $W = W(x,t)$ by

$$W(x,t)dx = P(x < w(t) \leq x + dx), \tag{3.20}$$

where $w(t)$ is a Brownian motion. $W(x,t)$ is the probability density function of the Brownian motion $u(t)$ at the fixed moment t. As we know,

$$W(x,t) = \frac{e^{-x^2/2t}}{\sqrt{2\pi t}}.$$

The graphs of $W(x,t)$ for several values of t are shown in Figure 3.

We see that the graphs become lower and wider as t increases. The increments of Brownian motion are independent. This means that if we know that at time t $w(t)$ is at x, then where it is at $t + \Delta t$ does not depend on where it was prior to the moment t. The relation between $W(x,t)$ and $W(x, t+\Delta t)$ is given by the Chapman-Kolmogorov equation

$$W(x, t+\Delta t) = \int_{-\infty}^{\infty} W(x+x', t)\Psi(x, x', \Delta t)dx', \tag{3.21}$$

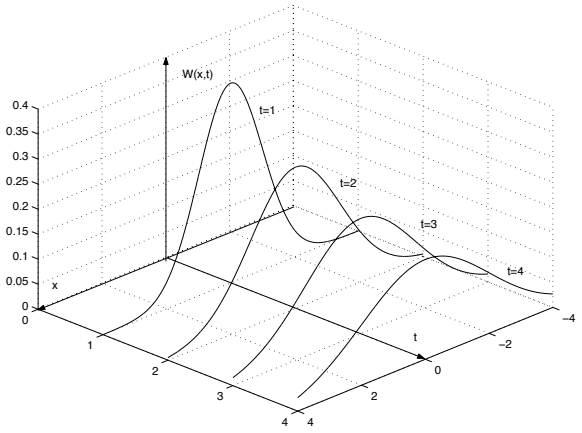

FIGURE 3. $W(x,t)$ for Brownian motion.

where

$$\Psi(x,x',\Delta t)=\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}} \tag{3.22}$$

is the "transition kernel." This equation states that the probability of reaching x at time $t+\Delta t$ is the sum of the probabilities of reaching x' at time t multiplied by the probability of going from x' to x during the time interval Δt.

Expand $W(x+x')$ in a Taylor series in x':

$$W(x+x')=W(x)+x'W_x(x)+\frac{(x')^2}{2}W_{xx}(x)+\frac{(x')^3}{6}W_{xxx}(x)+O((x')^4)$$

and substitute it into (3.21)

$$\begin{aligned}\int_{-\infty}^{\infty}W(x+x',t)\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}}dx' &= \int_{-\infty}^{\infty}W(x,t)\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}}dx'\\ &+\int_{-\infty}^{\infty}x'W_x(x,t)\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}}dx'+\frac{1}{2}\int_{-\infty}^{\infty}(x')^2W_{xx}(x,t)\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}}dx'\\ &+\frac{1}{6}\int_{-\infty}^{\infty}(x')^3W_{xxx}(x)\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}}dx'+\int_{-\infty}^{\infty}O((x')^4)\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}}dx',\end{aligned}$$

thus

$$\int_{-\infty}^{\infty}W(x+x',t)\frac{e^{-(x')^2/2\Delta t}}{\sqrt{2\pi\Delta t}}dx'=W(x)+0+\frac{W_{xx}(x)\Delta t}{2}+0+O(\Delta t^2).$$

Hence we have

$$W(x,t+\Delta t)=W(x,t)+\frac{\Delta t}{2}W_{xx}(x,t)+O(\Delta t^2).$$

Dividing by Δt we obtain

$$\frac{W(x,t+\Delta t)-W(x,t)}{\Delta t}=\frac{1}{2}W_{xx}(x,t)+O(\Delta t).$$

Letting $\Delta t\to 0$ we find:

$$\frac{\partial W}{\partial t}=\frac{1}{2}\,\frac{\partial^2 W}{\partial x^2}.$$

This is a "Fokker-Planck equation"- an equation that describes the time evolution of a one-time probability density for a stochastic process. We see that the Fokker-Planck equation for BM is the heat equation. This observation generalizes the relation between BM and the heat equation already described above.

A stochastic process is called a Markov process if what happens after time t is independent of what happened before time t, i.e., if $t'>t$ then

$$E[u(\omega,t')|u(\omega,t)]=E[u(\omega,t')|u(\omega,s),s\le t].$$

In other words, if we know $u(\omega,t)$ then knowing in addition $u(\omega,s)$ for $s<t$ does not help us to predict $u(\omega,t')$ for $t'>t$.

As discussed above, if $P(x<u(t)\le x+dx)=W(x,t)dx$, then $W(x,t)$ satisfies the Chapman-Kolmogorov equation

$$W(x,t+\Delta t)=\int W(x+x',t)\Psi(x',x,\Delta t)dx'$$

where Ψ is the "transition probability" from a state $x+x'$ at time t to the state x at time $t+\Delta t$. For a Markov process the transition probability does not depend on $W(x,s)$ for $s<t$. Brownian motion is by construction a Markov process because it has independent increments.

8. First Discussion of the Langevin Equation

Let $u(t,\omega)$ be a stochastic process defined by the following (formal) equation

$$\frac{du}{dt}=-au+\frac{dw}{dt}$$

where a is a positive constant and dw/dt is white noise, the derivative of a Brownian motion w. We know that this derivative does not exist in the classical sense thus the equation makes sense only formally (or else in the sense of distributions). A more sensible way to write the Langevin equation is

$$du=-audt+dw \tag{3.23}$$

where dw is the increment of Brownian motion. The meaning of (3.23) is defined by integrating from 0 to t

$$u(t) - u(0) = -a \int_0^t u dt + \int_0^t dw = -a \int_0^t u dt + w(t).$$

This is the Langevin equation (also known in some mathematical circles as the Ornstein-Uhlenback equation). It is an instance of a stochastic differential equation. The equation contains a term that is a random function of t, and the solution $u = u(\omega, t)$ should also be a random function of t that satisfies the equation for every ω in the probability space on which the equation is defined. The solution of this equation is known to mathematicians as the Ornstein-Uhlenbeck process.

If we omit the noise term in this equation and retain only the "damping" term $-aw$, the solution is a constant times e^{-at}, a pure decay. If on the other hand we keep the noise term but set $a = 0$ the solution of the equation is Brownian motion. In physics this equation is used to model the motion of a heavy particle under bombardment by lighter particles; the collisions with the lighter particles provide random instantaneous bursts of added momentum while the mean effect of the collisions is to slow the heavy particle down. We shall see in Chapter 6 that when this equation is used as a physical model the coefficient a, as well as the coefficient of the noise term which we have, rather arbitrarily, set equal to 1, acquire a direct physical meaning. The solution of this equation, with the coefficients interpreted correctly, is what physicists call Brownian motion.

Similarly to what we did in the case of Brownian motion, we want to find the equation satisfied by the probability density function of u, i.e., the Fokker-Planck equation for this problem. We choose an approximation for (3.23): integrating from nk to $(n+1)k$ where k is the timestep we have

$$u^{n+1} - u^n = -aku^n + w^{n+1} + w^n. \tag{3.24}$$

We choose k small enough so that $ak < 1$. The choice to evaluate the term $-aku$ at time nk is not just an arbitrary choice of approximation scheme but is dictated by the physics of the problem: we are constructing the solution step-by-step in time; what we have to work with when we go from time $t = nk$ to time $t = (n+1)k$ is the value of u we have previously calculated and the sample of BM in that time interval, and this is what we must use. The quantity $w^{n+1} - w^n$ in (3.24) is an increment of Brownian motion, therefore it is a Gaussian variable with mean zero and variance k. Equation (3.24) says that $u^{n+1} - u^n + aku^n$ is a Gaussian variable with mean zero and variance k. If u^n is known

then $P(x < u^{n+1} \leq x + dx)$ is

$$P(x < u^{n+1} \leq x + dx) = \frac{\exp\left(-\frac{(x-u^n+aku^n)^2}{2k}\right)}{\sqrt{2\pi k}}dx. \tag{3.25}$$

Since u^n is known this is exactly the transition probability from the point u^n at time nk to the point x at time $(n+1)k$. If we write $u^n = x + y$ then the Chapman-Kolmogorov equation is

$$W(x,(n+1)k) = \int_{-\infty}^{+\infty} W(x+y,nk)\Psi(x+y,x,k)dy.$$

Replacing Ψ by the expression we have just derived gives

$$W(x,(n+1)k) = \int_{-\infty}^{+\infty} W(x+y,nk)\frac{\exp\left(-\frac{(-y+ak(x+y))^2}{2k}\right)}{\sqrt{2\pi k}}dy.$$

After rearranging the exponent in the above we have

$$W(x,t+k) = \int_{-\infty}^{+\infty} W(x+y,t)\frac{\exp\left(-\frac{((1-ak)y-akx)^2}{2k}\right)}{\sqrt{2\pi k}}dy \tag{3.26}$$

where $t = nk$. The next step is to expand $W(x+y,t)$ around x. Up to fourth order we have

$$W(x+y,t) = W(x,t)+yW_x(x,t)+\frac{y^2}{2}W_{xx}(x,t)+\frac{y^3}{6}W_{xxx}(x,t)+O(y^4). \tag{3.27}$$

The expansion of $W(x+y,t)$ is substituted in (3.26) and we evaluate the different integrals that appear. Consider

$$I_1 = \int_{-\infty}^{+\infty} W(x,t)\frac{\exp\left(-\frac{((1-ak)y-akx)^2}{2k}\right)}{\sqrt{2\pi k}}dy.$$

To evaluate I_1 we make the change of variables $z = (1-ak)y$ and find

$$\begin{aligned} I_1 &= W(x)\int_{-\infty}^{+\infty}\frac{\exp\left(-\frac{(z-akx)^2}{2k}\right)}{\sqrt{2\pi k}}\frac{dz}{1-ak} \\ &= \frac{W(x)}{1-ak}\int_{-\infty}^{+\infty}\frac{\exp\left(-\frac{(z-akx)^2}{2k}\right)}{\sqrt{2\pi k}}dz \\ &= \frac{W(x)}{1-ak} \\ &= W(x)(1+ak+O(k^2)) \\ &= W(x)(1+ak)+O(k^2). \end{aligned} \tag{3.28}$$

The second integral is

$$I_2 = \int_{-\infty}^{+\infty} yW_x(x,t)\frac{\exp\left(-\frac{((1-ak)y-akx)^2}{2k}\right)}{\sqrt{2\pi k}}dy.$$

With the same change of variables we get:

$$\begin{aligned} I_2 &= W_x(x,t)\int_{-\infty}^{+\infty}\frac{z}{1-ak}\,\frac{\exp\left(-\frac{(z-akx)^2}{2k}\right)}{\sqrt{2\pi k}}\,\frac{dz}{1-ak} \\ &= \frac{W_x(x,t)}{(1-ak)^2}akx \\ &= W_x(x,t)(1+2ak+O(k^2))akx \\ &= W_x(x,t)akx + O(k^2). \end{aligned} \tag{3.29}$$

The third integral is

$$I_3 = \int_{-\infty}^{+\infty}\frac{y^2}{2}\,W_{xx}(x,t)\,\frac{\exp\left(-\frac{((1-ak)y-akx)^2}{2k}\right)}{\sqrt{2\pi k}}dy. \tag{3.30}$$

The same change of variables gives

$$\begin{aligned} I_3 &= W_{xx}(x,t)\int_{-\infty}^{+\infty}\frac{z^2}{2(1-ak)^2}\,\frac{\exp\left(-\frac{(z-akx)^2}{2k}\right)}{\sqrt{2\pi k}}\,\frac{dz}{1-ak} \\ &= W_{xx}(x,t)\frac{1}{2(1-ak)^3}(k+(akx)^2) \\ &= W_{xx}(x,t)\frac{k}{2}+O(k^2). \end{aligned} \tag{3.31}$$

The fourth integral is

$$I_4 = \int_{-\infty}^{+\infty}\frac{y^3}{6}W_{xxx}(x,t)\frac{\exp\left(-\frac{((1-ak)y-akx)^2}{2k}\right)}{\sqrt{2\pi k}}dy \tag{3.32}$$

which becomes

$$\begin{aligned} I_4 &= W_{xxx}(x,t)\int_{-\infty}^{+\infty}\frac{z^3}{6(1-ak)^3}\,\frac{\exp\left(-\frac{(z-akx)^2}{2k}\right)}{\sqrt{2\pi k}}\,\frac{dz}{1-ak} \\ &= W_{xxx}(x,t)\,\frac{1}{6(1-ak)^4}\,(3axk^2+(akx)^3) \\ &= W_{xxx}(x,t)O(k^2). \end{aligned} \tag{3.33}$$

As we see the fourth integral contributes only terms of order $O(k^2)$ and higher; the same is true of the next terms in the expansion which

have been omitted. Collecting (3.26), (3.27), (3.28), (3.29), (3.31), and (3.33) and we find

$$W(x,t+k) = W(x,t) + W(x,t)ak + W_x(x,t)akx + \frac{k}{2}W_{xx}(x,t) + O(k^2),$$

and

$$\frac{W(x,t+k) - W(x,t)}{k} = W(x,t)a + W_x(x,t)ax + \frac{1}{2}W_{xx}(x,t) + O(k),$$

and finally, as we make $k \to 0$,

$$W_t(x,t) = (axW(x,t))_x + \frac{1}{2}W_{xx}(x,t).$$

This is the Fokker-Planck equation corresponding to the solution of the Langevin equation (3.23).

Given a Markovian stochastic process we can construct its Fokker-Planck equation and vice-versa. An interesting pair of a stochastic ordinary differential equation and the corresponding Fokker-Planck equation arises in two-dimensional incompressible fluid mechanics. If we consider a fluid having velocity $\mathbf{u} = (u, v)$ and vorticity $\xi = v_x - u_y$, where (x, y) represents a point in physical space, then the equation for the evolution of the vorticity is

$$\frac{\partial \xi}{\partial t} + (\mathbf{u} \cdot \nabla)\xi = \frac{1}{Re}\Delta\xi \tag{3.34}$$

where Re is the Reynolds number of the flow. If we assume that $\xi \geq 0$ and $\int \xi = 1$ then (3.34) is the Fokker Planck equation of the following system of stochastic ordinary differential equations

$$d\mathbf{x} = \mathbf{u}dt + \sqrt{\frac{2}{Re}}d\mathbf{W}.$$

Here $\mathbf{x}$ is the position of the point where the vorticity is ξ, and $\mathbf{W}$ is a two-dimensional Brownian motion. Each of these particles carries a fixed amount of vorticity and the corresponding evolution of the density solves the vorticity partial differential equation. There is one equation per point in the support of ξ (i.e., for every point (x, y) such that $\xi(x, y) \neq 0$). The velocity $\mathbf{u}$ depends on the whole vorticity field at each instant t, so this equation is nonlinear and couples the BM's that correspond to different points in physical space, as one should expect given that the original equation of motion is nonlinear.

9. Solution of a Nonlinear Differential Equation by Branching Brownian Motion

So far, with the exception of the short comments at the end of the previous section, all the equations we have been solving have been linear. Now we give an example of how a variant of BM can be used to solve a nonlinear partial differential equation. The equation we work with is the Kolmogorov-Petrovski-Piskunov (KPP) equation,

$$v_t - v_{xx} = v^2 - v,$$

for which we prescribe initial data $v(x, t = 0) = \phi(x)$. This equation is an important model in combustion theory and in biology. We are looking, as before, for a representation of the solution v at a point (x, t).

Start a BM w going backwards in time from (x, t) and let it run until a time $t - t_1$ drawn at random from the exponential distribution, $P(y < t_1 \leq y+dy) = \exp(-y)dy$. Start two independent BM's running backwards from $(w(t_1), t_1)$, call them w_{11}, w_{12}, until new exponentially distributed times $t - t_{11}, t - t_{12}$. At each stopping time split the branch of the BM into two independent BM's. If the time becomes negative for any branch, stop. The result is a backward tree with roots that cross the x-axis. Let the intersections of the tree with the x-axis be $x_1, x_2, \ldots, x_n$, $n \geq 1$, and associate with the tree the product of initial values $\Xi = \phi(x_1)\phi(x_2)\cdots\phi(x_n)$; the claim is that the expected value of this product is the solution we want:

$$v(x, t) = E[\Xi] = E[\phi(x_1)\cdots\phi(x_n)].$$

We take this opportunity to introduce a notation which will be widely used in Chapter 6. Let Δ be the second derivative operator in the space variable x: $\Delta f = f_{xx}$ for a smooth function f. Just as the solution of the equation $v' - av = 0, v(0) = v_0$, a =constant, is $e^{at}v_0$, we will symbolically write the solution of the heat equation $v_t - \Delta v = 0, v(x, 0) = \phi$, as $v(t) = e^{t\Delta}\phi$ (this is the "semigroup" notation). For $v(x, t)$, which is the function $v(t)$ evaluated at x, we write $v(x, t) = (e^{t\Delta}\phi)(x)$. We know that $(e^{t\Delta}\phi)(x) = E[\phi(x) + w(t)]$, where as before w is BM. One can readily understand the identity $e^{(t+s)\Delta} = e^{t\Delta}e^{s\Delta}$ and check its validity (this is the "semigroup property").

The probability that the first branching occurs at a time t_1 larger than t is $\int_t^\infty e^{-s}ds = e^{-t}$; if this happens the number Ξ attached to the tree is $\phi(x) + w(t)$ whose expected value, conditioned by $t_1 > t$, is $(e^{t\Delta}\phi)(x)$. Supppose to the contrary that t_1 occurs in a time interval $(s, s + ds)$ earlier than t (this happens with probability $e^{-s}ds$ by construction). Then the two branches of the tree that start from

the point $(t_1, w(t_1))$ carry two independent samples of the Ξ for the subtrees that start from that point; the expected value of their product is $v^2(w(t_1), t-s)$. Averaging over all the BM's that start from (x,t) and branch at a time between s and $s+ds$ gives $(e^{s\Delta}v^2(t-s))(x)$.

Collecting all these facts we find:

$$\begin{aligned} v(x,t) = E[\Xi] &= e^{-t}e^{t\Delta}\phi + \int_0^t e^{-s}e^{s\Delta}v^2(t-s)ds \\ &= e^{-t}e^{t\Delta}\phi + \int_0^t e^{s-t}e^{(s-t)\Delta}v^2(s)ds, \end{aligned}$$

where the last identity is obtained by the change of variables $s' = t-s$ and the prime is dropped. Differentiate this expression with respect to t, noting that $\Delta e^{-t} = e^{-t}\Delta$ (differentiation with respect to x and multiplication by a function of t commute), and find that v satisifies the equation; it is obvious that $v(x,0) = \phi$.

10. A Brief Introduction to Stochastic ODEs

We have solved above a particular stochastic differential equation—the Langevin equation; we now make some comments about more general stochastic ordinary differential equations (SODE's) of the form

$$du = a(t,u(t))dt + b(t,u(t))dw, \tag{3.35}$$

where w is Brownian motion. The meaning of this equation is defined by

$$u(t) - u(0) = \int_0^t a(s,u(s))ds + \int_0^t b(s,u(s))dw.$$

The first integral is well-defined while, as we shall now see, the second is not. Integrals of this form are called stochastic integrals. Let us figure out in what sense we can understand them.

Let $f(t)$ be a function defined on an interval $[a,b]$. A partition of $[a,b]$ is a set of points $\{t_i\}_{i=0}^n$ such that

$$a = t_0 < t_1 < t_2 < \ldots < t_n = b.$$

Definition. The variation of $f(t)$ on $[a,b]$ is defined by

$$\mathrm{Var}(f(t)) = \sup_{\text{all partitions}} \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|. \tag{3.36}$$

If the sup is finite f is said to have bounded variation; Brownian motion does not have bounded variation. Stieltjes integrals of the form

$\int g(t)df(t)$ make sense only when the increment function f has bounded variation and therefore

$$\int_0^t b(s, u(s))dw$$

is not well-defined as a Stieltjes integral.

The way to make sense of the stochastic integrals is to approximate $b(t, u(s))$ by a piecewise constant function, i.e.,

$$\int_0^t b(s, u(s))dw \approx \sum_{i=0}^{n-1} b_i du_i = \sum_{i=0}^{n-1} b_i(w(t_{i+1}) - w(t_i)),$$

where $\{t_i\}_{i=0}^n$ is a partition of $[0, t]$, and then consider the limits of the sum as one makes the largest interval $t_i - t_{i-1}$ in the partition go to zero. Now one has to decide how to pick the b_i's. There are two common choices:

(1) The b_i's are evaluated at the left ends of the intervals, i.e.,

$$b_i = b(t_i, u(t_i)).$$

(2) The b_i's are the average of the endpoints

$$b_i = \frac{1}{2}\left[b(t_i, u(t_i)) + b(t_{i+1}, u(t_{i+1}))\right].$$

Choice 1 defines the Ito stochastic integral while choice 2 defines the Stratonovich stochastic integral.

EXAMPLE. Suppose $b(t, u(t)) = w(t)$. (To apply the analysis to the SODE above we could assume that the SODE has the solution $u(t) = w(t)$, however, all we are trying to show is that the two definitions of the integrals give different answers.) Then in the Ito case:

$$I_1 = \int_0^t w dw \approx \sum_{i=0}^{n-1} w(t_i)(w(t_{i+1}) - w(t_i)).$$

This is of course a random variable; the expected value of this random variable is zero, as one can see from the properties of BM:

$$E[I_1] = 0.$$

In the Stratonovich case we find for the stochastic integral:

$$\begin{aligned} I_2 &= \int_0^t u du \approx \sum_{i=0}^{n-1} \frac{1}{2}(w(t_{i+1}) + w(t_i))(w(t_{i+1}) - w(t_i)) \\ &= \sum_{i=0}^{n-1} \frac{1}{2}(w^2(t_{i+1}) - w^2(t_i)) \\ &= \frac{1}{2}\left[w^2(t_1) - w^2(t_0) + w^2(t_2) - w^2(t_1) + \ldots + w^2(t_n) - w^2(t_{n-1})\right] \\ &= \frac{1}{2}\left[w^2(t_n) - w^2(t_0)\right] = \frac{1}{2}w^2(t), \end{aligned}$$

and the expected value of this integral is

$$E[I_2] = \frac{t}{2}.$$

The fact that the expected values of the two integrals are so different is of course enough to show that the integrals themselves are different. This is very different from the situation in ordinary calculus, where the value of an integral is independent of the choice of points in the Riemann sums. How the stochastic integral is defined makes a big difference to the meaning of a stochastic differential equation. For the sake of definiteness we shall assume henceforth, when this makes a difference, that we are dealing with stochastic differential equations in the sense of Ito. When b in (3.35) is a constant (as has been the case so far in these notes) there is no ambiguity.

11. References

1. L. Arnold, *Stochastic Differential Equations*, Wiley, NY, (1973)
2. R. Bhattacharya, L. Chen, S. Dobson, R. Guenther, C. Orum, M. Ossiander, E. Thomann, and E. Waymire, *Majorizing kernels and stochastic cascades with applications to incompressible Navier-Stokes equations*, Trans. Am. Math. Soc. 355 (2003), pp. 5003-5040.
3. S. Chandrasekhar, *Stochastic problems in physics and astronomy*, Rev. Modern Phys. 15, (1943), pp. 1-88, reprinted in N. Wax, Selected Papers on Noise and Stochastic Processes, Dover, NY, (1954).
4. A.J. Chorin, *Numerical study of slightly viscous flow*, J. Fluid Mech. 57 (1973), pp. 785-796.
5. A.J. Chorin, *Vortex methods*, in Les Houches Summer School of Theoretical Physics, 59, (1995), pp. 67-109.

6. A.J. Chorin, *Accurate evaluation of Wiener integrals*, Math. Comp. 27 (1973), pp. 1-15.
7. R. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw Hill, NY, (1965)
8. K. Ito and H. McKean, *Diffusion Processes and their Sample Paths*, Springer, Berlin, 1974.
9. M. Kac, *Probability and Related Topics in the Physical Sciences*, Interscience Publishers, London, (1959).
10. P. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer, Berlin, 1992.
11. J. Lamperti, *Probability*, Benjamin, NY, 1966.
12. H. McKean, *Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov*, Comm. Pure Appl. Math 27, (1975), pp. 323-331.
13. S. Neftci, *An Introduction to the Mathematics of Financial Derivatives*, Academic, New York, 2000.
14. R. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, Colloquium publications, American Math. Soc. vol. 19, Providence, RI, (1934).
15. L. Schulman, *Techniques and Applications of Path Integration*, Wiley, NY, 1981.
16. N. Wiener, *Nonlinear Problems in Random Theory*, MIT Press, Cambridge Mass., (1958).